Chasing Infinitesimals

Francesco Lucente Stabile
Senior Seminar - MAT 490
Relator: *Dr. Pedro Poitevin*
Mathematics Department
Salem State University

July 26, 2019

The smallest ghosts

The ability to see smaller and smaller objects has always been one of the hardest challenges of humankind together with the ability to see bigger and bigger things.

This ability grew together with a general advancement in technology and progress that built newer and more powerful tools, such as microscopes and telescopes. Already in the fifth century before Christ Democritus started to talk about indivisible entities (atoms) that made up the entire universe. The idea of a smallest part making up the whole seems to be the foundation of science and logic, and progress so far did not disappoint our early hypothesis. In fact today with the particles accelerators we are able to detect subatomic particles and the radio telescopes help us to see enormous objects like the further Quasars.

The journey chasing infinitesimals and infinite numbers in mathematics shares many analogies with the one in science. Infinitesimals have been theorized since the beginning of logic. As far as math can remember infinitely small objects and distances where known to exist, but they where no more then ghosts. Every mathematicians knew they where there, they felt them, but no one could see them, since the tools to detect such small entities was still far to be discovered.

Archimedes started to see geometric objects as made up of an infinite number of infinitely small elements. Eudoxus' reductio ad absurdum was a rudimentary step towards the more advanced calculus that was used by Archimedes to find the area and volumes of many solids. Of course the method involved slicing the solids in laminae of infinitely small thickness and add them up together. Archimedes himself said about the method: "This has not therefore been proved, but a certain impression has been created that the conclusion is true." This amounted to suggesting that he knew that mathematical objects are made by smallest parts but at that time this could be no more than a good intuition.

Leibniz refined the idea of the reductio ad absurdum discovering the differential and integral calculus. It was the seventeenth century when he introduced the symbol dx to represent an infinitely small quantity, but as he himself wrote, "... as small as you please, so that the error that one may assign may be less that a certain assign quantity". The infinitesimals whispered to Leibniz too, but since it did not have a fishnet that was thin enough to catch them, he had the brilliant idea of defining the dx as a quantity that can be omitted since it can be chosen as small as we need to omit it. But even then, no infinitesimals where clearly seen. Notably, Leibniz's conception of infinitesimal differs philosophically ever so slightly from Newton's (in ways that we won't get into in this work), and I have been told Leibniz had things more right than Newton.

Even Euler, was chasing these little ghosts. He just knew they existed and they acted like normal numbers. For him it existed a number ω that was so small that $a^{\omega} = 1+\psi$, where also ψ was infinitely small. However in his works, he did not explored the infinitesimals as a measurable quantity to be known exactly, but he instead gave Leibniz and Newton's work for granted and took advantage of the concept that some numbers are as small as we can make them. It was around the 1750.

Then, in 1966, three hundreds years later Robinson finally saw them. He caught the infinitesimals. Doing so required deep logical insight. Later on, Luxembourg came up with tools to finally see them, like the man who invented the microscope. It was the moment were humans were given the maximum magnification machinery.

We will now explore a combination of both approaches to build a numerical system where the infinitesimals lay and can be seen in all their splendor.

Why couldn't we see the infinitesimal before 1966?

Because we were looking in the wrong place. The most advanced, dense field at that time was the real numbers. The problem mathematicians had with the "magnification" was the way that $\mathbb{R}$ is built.

Despite $\mathbb{R}$ being a miracle field in which all irrational and transcendental numbers are included, and despite also the extreme density of the real number line, infinitesimals cannot be spotted on it (even though the lay just right around each corner), because of the way the real numbers are defined and "equalized".

According to the classical construction of the real numbers as equivalence classes of Cauchy sequences of rational numbers, if $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are Cauchy sequences of rational numbers, then $\lim_{n\to\infty} a_n$ and $\lim_{n\to\infty} b_n$ are real numbers, and these real numbers are the same if and only if the sequence $(a_n - b_n)_{n=1}^{\infty}$ converges to 0. As it turns out, the equivalence relation on $\mathbb{R}$ that is above defined is guaranteed to hide the infinitesimals.

Consider the Cauchy sequence $\left(\frac{1}{n}\right)_{n=1}^{\infty}$. This sequence is an obvious candidate for giving rise to an infinitesimal, since intuitively the terms decrease infinitely as the denominator gets large without limits, but they never become zero. However, given the classical equivalence relation used to build the reals as equivalence classes of Cauchy sequences of rational numbers, this sequence is simply the real number 0.

Reiterating, if we chose an arbitrary $\epsilon > 0$ and a number N such that $N > \frac{1}{\epsilon}$, then for every $n \geq N$, we can see that $|\frac{1}{n} - 0| \leq \frac{1}{n} \leq \frac{1}{N} < \epsilon$. This means that $\left(\frac{1}{n}\right)_{n=1}^{\infty}$ is zero, or better it is of the equivalence class of zero. So basically it represents zero together with other "ghosts" like $\left(\frac{1}{n^2}\right)_{n=1}^{\infty}$, $\left(\frac{1}{n^3}\right)_{n=1}^{\infty}$ and so on, none excluded. All the Cauchy sequences that are eventually ϵ- close to 0, are 0. So we are left with no hint of an infinitesimal.

The thing is we all see that they are not zero, but the real number system is entirely built on the concept of ϵ- closeness, so if we deny them to be zero, the consequence is simply the disappearing of the real number line. However giving up on the idea that $\left(\frac{1}{n}\right)_{n=1}^{\infty}$ is an infinitesimal seems like giving up on an obvious truth, or at the very least an obvious path for enlightenment.

Then let's start with analyzing sequences of rationals in a different way and build a field where the equivalent relation is favorable to our intuition. A good start is to observe that the sequences $\left\langle 1, \frac{1}{2}, \frac{1}{3}, \ldots \right\rangle$ and $\langle 0, 0, 0, \ldots \rangle$ have nothing in common if compared component wise. Or we can say, anticipating the terminology that we will use from now on, that the terms do not agree.

The first hurdle we find is that we cannot define an equivalence relation between two sequences limited on a total agreement terms-wise, or else $\langle 0, 0, 0, \ldots \rangle$ and $\langle 7, 0, 0, 0, \ldots \rangle$ would not be equal and we would have big issues in defining what the first sequence actually represents if not 0.

Then we introduce a new amazing idea, simple at first but that does not come without complications: Let two sequences be equivalent if they agree on a *large* number of terms. (We leave the concept of "large" vague for the moment, but we'll get back to it soon.)

Specifically, let $r = \langle r_1, r_2, r_3, \ldots \rangle$ and $s = \langle s_1, s_2, s_3, \ldots \rangle$ be two sequences, so that

$$r \equiv s \text{ if and only if the set } E_{r,s} = \{n | r_n = s_n\} \text{ is large.}$$

This would immediately set $\langle \frac{1}{n} \rangle \neq 0$ since $E_{\langle \frac{1}{n} \rangle, 0} = \emptyset$, so obviously not large. (We will need to define our concept of largeness to reflect the intuition that $\emptyset$ is not large, of course.)

However we encounter now the first high hurdle. We can tell (intuitively) that $\emptyset$ is not large but, what's large?

It is hard to start with showing a large object, as we will be able to do in the future, but we can instead begin by defining the properties of largeness, so that we can recognize large objects by their properties, and we can build tools that are able to recognize and catch objects with such properties.

How largess look like

The first property of largeness comes from the fact that we want to ensure any sequence to be equal to itself, so we want (or we need) the set of the natural numbers to be large. This is because we need $E_{r,r} = \{n | r_n = r_n\} = \mathbb{N}$ to be sufficiently large for the equivalency.

Then, since the equivalence relation must be transitive, if A and B are large sets, then we hope $A \cap B$ also to be large. This property seems wanting to exclude many intuitively large sets, but if E_{rs} and E_{st} are large, then $E_{rt} = \{n | r_n = t_n\}$ is also large by the transitivity. But $E_{rs} \cap E_{st} = E_{rt}$. So more in general if $A \cap B \subseteq C$, then C is large. And this also implies that every superset of a large set is large.

Moreover, for many reasons, we can say that the empty set cannot be large. In fact, if $\emptyset$ is large, then every superset of $\emptyset$ is large, so every subset of $\mathbb{N}$ would be large. Also if this property wouldn't hold, then not even the first properties would hold.
Finally, as consequence of the previous properties we have that for every subset A of $\mathbb{N}$ either A or A^c are large. In fact, for $A^c = \mathbb{N} - A$, we have that $A \cap A^c = \emptyset$. Since we established that

the empty set is not large, then by the first property, A and A^c cannot be large simultaneously.
Summarizing:

1. $\mathbb{N}$ is large.

2. $\emptyset$ is not large.

3. If A and B are large sets and $A \cap B \subseteq C$, then C is large.

4. For $A \subseteq \mathbb{N}$, A or A^c is large.

With this properties given to largeness, we can now build an instrument that help us to collect objects with such properties.

As a mechanical tool that filters small things are retains the big ones in its thin or rade net, we call these objects **filters**.

Filters and Ultrafilters

A filter denoted by this elegant, cursive $\mathcal{F}$, is a collection of sets, with the characteristic that inside the filter only large sets get stuck. Since we do not know yet what a large thing is, but we know some properties of largeness, we will see how only objects with those properties will be found in the filters, so in fact how the properties of the subsets of $\mathcal{F}$ mirror the properties of large objects.

If I is a non empty set, and $\mathcal{P}(I) = \{A : A \subseteq I\}$ is the power set of I, meaning the set containing all the possible subsets of I, then $\mathcal{F}$ is a non empty subset of $\mathcal{P}(I)$, with the following properties:

1. If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.

   The intersection property will ensure the transitivity via the agreement sets that will get stuck on the filter.

2. if $A \in \mathcal{F}$ and $A \subseteq B \subseteq I$, then $B \in \mathcal{F}$.

The superset property, consequence of the intersection property, will help to determine whether a set B is contained in a filter, simply by checking that

$$A_1 \cap A_2 \cap \cdots \cap A_n \subseteq B$$

for some n and some $A_i \in \mathcal{F}$.

Consequently a filter that contains the empty set is indeed the power set, since all the subsets of I are supersets of $\emptyset$ and they are in the filter. Moreover, for every set I, $I \in \mathcal{F}$. Note that the set containing I is the smallest filter on I. In fact, $I \cap I \subseteq I \in \mathcal{F}$, and $I \subseteq I$ and it is indeed the only supersets of I. So I has the properties to get stuck in the filter.

An ultrafilter $\mathcal{U}$ is a filter that satisfies:

3. For any $A \subseteq I$, either $A \in \mathcal{U}$ or $A^c \in \mathcal{U}$.

We will see that an ultrafilter is a proper filter (so $\emptyset \notin \mathcal{F}$, and $\mathcal{F} \neq \mathcal{P}(I)$) that cannot be extended to a larger proper filter.

There exist different types of filters, with various properties and characteristics, constructed to catch different objects, and we need to be familiar with some of them, which will be the ones that will help us to construct the hyperreal number system.

The Principal Ultrafilter generated by i is the set of all the subsets of a nonempty set I that contains i. In notation $\mathcal{F}^i = \{A \subseteq I | i \in A\}$.

The filter generated by $\mathcal{H}$, for $\emptyset \neq \mathcal{H} \subseteq \mathcal{P}(I)$, is the smallest filter on I including $\mathcal{H}$. Which is the collection $\mathcal{F}^{\mathcal{H}} = \{A \subseteq I | B_1 \cap \cdots \cap B_n \subseteq A \text{ for some } n \text{ and some } B_i \in \mathcal{H}\}$.
$\mathcal{F}^i$ is a special case of the filter generated by $\mathcal{H}$, that is the case

where $\mathcal{H}$ has only one member $\{i\}$.

The cofinite filter, or Frechet filter (which is not an ultrafilter but it is contained in an ultrafilter) is the collection of all the subsets of I whose complement is finite.

A relevant fact to observe is that an ultrafilter that contains a finite set, then it also contains a one-element set, so it is principal. It follows that a non-principal ultrafilter contains all cofinite sets. (Fact 1)

Now that our tools are built let's make sure that the work in the way we expect when applied on sets we need to determine as large or not. We will need to see if there is indeed an ultrafilter on $\mathbb{N}$.

First note that Zorn's Lemma (the set theory version of the axiom of choice) states that if $(P, \leq)$ is a poset (partially ordered set), and every linearly ordered ascending chain on P has an upperbound, then P contains a $\leq$-maximal element.

A non rigorous proof of this lemma is given by considering the order of $\mathbb{N}$. Since in fact $|\mathbb{N}| = \omega$, then each ascending chain in P has a maximal capacity of ω elements before it can still be considered a set. So, since each chain has an upperbound ω, then, P has a maximal element.

Then we need to check that $(P, \subseteq)$ is a poset, so that $\subseteq$ is a relation on P.

So let $A \subseteq P$. Then $A \subseteq A$. Then $\subseteq$ is reflexive.

Second let $A, B \subseteq P$. Then if $A \subseteq B$ and $B \subseteq A$, then $A = B$. hence $\subseteq$ is antisymmetyric.

Finally if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$. Thus $\subseteq$ is also transitive and we have shown that $(P, \subseteq)$ is a poset.

$\square$

Next we need to show that every infinite set has a non principal ultrafilter on it:

So let I be an infinite set. Note that the Frechet Filter $\mathcal{F}^{co}$ is proper, and so it has the FIP. Hence it is included in an ultrafilter U.

Observe that $\forall i \in I$, $I - \{i\} \in \mathcal{F}^{co} \in \mathcal{U}$. Hence $\{i\} \notin \mathcal{U}$, but $\{i\} \in \mathcal{U}^i$. We conclude that $\mathcal{U} \neq \mathcal{U}^i$.
Hence we have shown that the ultrafilter on an infinite set is non principal.

$\square$

At this point, we can summarize and observe some facts that are fundamental for our further work: since $\mathbb{N}$ is infinite, then it has an ultrafilter on it and it is non principal. Then by (Fact 1) since the ultrafilter is non principal, then it contains all cofinite sets. (Fact 2).

Constructing a field

let $\mathbb{R}^{\mathbb{N}}$ be the set of all sequences of real numbers. Then the elements of $\mathbb{R}^{\mathbb{N}}$ are of the form $r =< r_1, r_2, r_3, ... >$.

Observe that, for $r, s \in \mathbb{R}^{\mathbb{N}}$

$$r \oplus s =< r_n + s_n : n \in \mathbb{N} >$$

and tat

$$r \odot s =< r_n \cdot s_n : n \in \mathbb{N} >$$

.

We can see that $< \mathbb{R}^{\mathbb{N}}, \oplus, \odot >$ is a ring but not a field, where $0 =< 0, 0, 0, ... >$ and $1 =< 1, 1, 1, ... >$. However, $< 1, 0, 1, 0, ... > \odot < 0, 1, 0, 1, ... >=< 0, 0, 0, ... >= 0$. Thus every sequence that contains at least one zero term have no multiplicative inverse.

Now we begin to get into the heart of the matter.

Let's use (Fact 2). Since $\mathbb{N}$ has a non principal ultrafilter on it, we define the equivalence relation $\equiv$ on $\mathbb{R}^{\mathbb{N}}$ as

$$r \equiv s \text{ iff } \{n \in \mathbb{N} : r_n = s_n\} \in \mathcal{U}.$$

We practically begun using the idea of largeness we built earlier as a help in defining the equivalence relation on $\mathbb{R}^{\mathbb{N}}$ this way: r is equivalent to s if the set containing the number of the terms that they agree on is large enough to be contained in an ultrafilter.

Then if $r \equiv s$ we say that r and s agree almost everywhere (modulo $\mathcal{U}$).

It is not ordinary to read a mathematical statement that is not either true or false, but whose truth is indeed a set of values, and just to underline and emphasize this concept we write

$$r \not\equiv s \text{ iff } \{n \in \mathbb{N} : r_n = s_n\} \notin \mathcal{U}.$$

Now note that the relation $\equiv$ can be extended also to $<, >, \leq$ and $\geq$ in $\mathbb{R}^{\mathbb{N}}$ in the following ways:

$$\begin{aligned} r < s &= \{n \in \mathbb{N} : r_n < s_n\} \\ r > s &= \{n \in \mathbb{N} : r_n > s_n\} \\ r \leq s &= \{n \in \mathbb{N} : r_n \leq s_n\} \\ r \geq s &= \{n \in \mathbb{N} : r_n \geq s_n\}. \end{aligned}$$

A new system

Let ${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}} / \sim \mathcal{U}$ be the set of equivalence classes of $\mathbb{R}^{\mathbb{N}}$ by $\equiv$.

Then for $r \in \mathbb{R}$, $[r] \in^* \mathbb{R}$, and if $[r]$ is the equivalence class of a sequence $r \in \mathbb{R}^{\mathbb{N}}$, then

$${}^*\mathbb{R} = \{[r] : r \in \mathbb{R}^{\mathbb{N}}\}.$$

Observe that $+, \cdot, <, >, \leq$ and $\geq$ are also well defined and

$$[r] < [s] \text{ iff } \{n \in \mathbb{N} : r_n < s_n\} \in \mathcal{U}.$$

It might not seen so obvious but we have finally built a machinery whose magnification is so powerful to finally show us infinitesimals in all their beauty!

So reconsider our candidate sequence $< \frac{1}{n} >$ as an element of $\mathbb{R}^{\mathbb{N}}$. Note that the set of the equivalence class of the sequence $[\frac{1}{n}]$ is an element of ${}^*\mathbb{R}$.

Now, the sequence $0 =< 0, 0, 0, ... >$ and $\frac{1}{n} =< 1, \frac{1}{2}, \frac{1}{3}, ... >$ agree everywhere (modulo $\mathcal{U}$) under $<$.

So $[0] < [\frac{1}{n}]$ since $\{n \in \mathbb{N} : 0 < \frac{1}{n}\} \in \mathcal{U}$.

So, again, we have that $[0] < [\frac{1}{n}]$.

Now let $\epsilon > 0 \in \mathbb{R}$, and let $\epsilon =< \epsilon, \epsilon, \epsilon, ... > \in \mathbb{R}^{\mathbb{N}}$. Then $[\epsilon] \in {}^*\mathbb{R}$.

Let us analyze the two sequences under the relation $<$.
Notice that $[\frac{1}{n}] < [\epsilon]$. In fact $[\epsilon]$ might be smaller, if $0 < \epsilon < 1$, but only for a finite amount of terms, then it will eventually be larger. Hence the terms whose index in in the set of agreement of $[\frac{1}{n}]$ being smaller than $[\epsilon]$ is cofinite! So by magic

$$[\frac{1}{n}] < [\epsilon] \text{ since } \{n \in \mathbb{N} : \frac{1}{n} < \epsilon \text{ is cofinite}\} \in \mathcal{F}^{co} \in \mathcal{U},$$

where $\mathcal{F}^{co}$ is a Frechet filter.

Thus we have seen that $[0] < [\frac{1}{n}] < \epsilon$, for all $\epsilon > 0$ and all $n \in \mathbb{N}$.
Here we just met an infinitesimal!

Bibliography

Goldblatt, Robert. *Lectures on Hyperreals: an Introduction to Nonstandard Analysis.* Springer, 1998

Nelson, Edward. *Radically Elementary Probability Theory.* Princeton University Press, 1987.

Luxemburg, W. A. *"What Is Nonstandard Analysis?" The American Mathematical Monthly*, vol. 80, no. 6, 1973, pp. 38–67. JSTOR, JSTOR, www.jstor.org/stable/3038221.